UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
10/686,459	10/15/2003	David Crawford Gibbon	113344Con-1	7802

26652 7590 05/30/2006

AT&T CORP.
ROOM 2A207
ONE AT&T WAY
BEDMINSTER, NJ 07921

EXAMINER	
LERNER, MARTIN	

ART UNIT	PAPER NUMBER
2626	

DATE MAILED: 05/30/2006

Please find below and/or attached an Office communication concerning this application or proceeding.

Office Action Summary	Application No.	Applicant(s)	
	10/686,459	GIBBON ET AL.	
	Examiner	Art Unit	
	Martin Lerner	2626	

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

1)☐ Responsive to communication(s) filed on _____.

2a)☐ This action is **FINAL**. 2b)☒ This action is non-final.

3)☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

4)☒ Claim(s) 1 to 25 is/are pending in the application.

4a) Of the above claim(s) _____ is/are withdrawn from consideration.

5)☐ Claim(s) _____ is/are allowed.

6)☒ Claim(s) 1 to 25 is/are rejected.

7)☐ Claim(s) _____ is/are objected to.

8)☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

9)☒ The specification is objected to by the Examiner.

10)☐ The drawing(s) filed on _____ is/are: a)☐ accepted or b)☐ objected to by the Examiner.

Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).

11)☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

12)☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

a)☐ All b)☐ Some * c)☐ None of:

1.☐ Certified copies of the priority documents have been received.

2.☐ Certified copies of the priority documents have been received in Application No. _____.

3.☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

1) ☒ Notice of References Cited (PTO-892)
2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)
3) ☒ Information Disclosure Statement(s) (PTO-1449 or PTO/SB/08) Paper No(s)/Mail Date _____.
4) ☐ Interview Summary (PTO-413) Paper No(s)/Mail Date. _____ .
5) ☐ Notice of Informal Patent Application (PTO-152)
6) ☐ Other: _____.

DETAILED ACTION

Specification

1. The disclosure is objected to because of the following informalities:

2. On page 1, lines 5 to 12, continuity data should be updated to reflect U.S. Patent No. 6,714,909, issued 30 May 2004; U.S. Patent No. 6,317,710, issued 13 November 2001; and U.S. Patent No. 6,801,895, issued 05 October 2004.

On page 11, line 5, "different" should be –difference—.

On page 11, line 8, "Ration" should be –Ratio—.

On page 21, line 1, "maybe" should be –may be—.

On page 22, line 11, "maybe" should be –may be—.

Appropriate correction is required.

Double Patenting

3. The nonstatutory double patenting rejection is based on a judicially created doctrine grounded in public policy (a policy reflected in the statute) so as to prevent the unjustified or improper timewise extension of the "right to exclude" granted by a patent and to prevent possible harassment by multiple assignees. A nonstatutory obviousness-type double patenting rejection is appropriate where the conflicting claims are not identical, but at least one examined application claim is not patentably distinct from the reference claim(s) because the examined application claim is either anticipated by, or would have been obvious over, the reference claim(s). See, e.g., *In re Berg*, 140 F.3d 1428, 46 USPQ2d 1226 (Fed. Cir. 1998); *In re Goodman*, 11 F.3d 1046, 29 USPQ2d 2010 (Fed. Cir. 1993); *In re Longi*, 759 F.2d 887, 225 USPQ 645 (Fed. Cir. 1985); *In re Van Ornum*, 686 F.2d 937, 214 USPQ 761 (CCPA 1982); *In re Vogel*, 422 F.2d 438, 164 USPQ 619 (CCPA 1970); and *In re Thorington*, 418 F.2d 528, 163 USPQ 644 (CCPA 1969).

A timely filed terminal disclaimer in compliance with 37 CFR 1.321(c) or 1.321(d) may be used to overcome an actual or provisional rejection based on a nonstatutory double patenting ground provided the conflicting application or patent either is shown to be commonly owned with this application, or claims an invention made as a result of activities undertaken within the scope of a joint research agreement.

Effective January 1, 1994, a registered attorney or agent of record may sign a terminal disclaimer. A terminal disclaimer signed by the assignee must fully comply with 37 CFR 3.73(b).

4. Claims 1, 5, 13, and 14 are rejected on the ground of nonstatutory obviousness-type double patenting as being unpatentable over claims 1 to 8 of U.S. Patent No. 6,714,909. Although the conflicting claims are not identical, they are not patentably distinct from each other because the corresponding claims set forth the same subject matter with respect to steps of separating a multimedia stream, segmenting the components, identifying a target speaker, identifying semantic boundaries, generating a summary, deriving a topic, and generating a multimedia description. Also, the corresponding claims set forth the same subject matter with respect to clip level features, and frame level features in three subbands.

Claim Rejections - 35 USC § 102

5. The following is a quotation of the appropriate paragraphs of 35 U.S.C. 102 that form the basis for the rejections under this section made in this Office action:

> A person shall be entitled to a patent unless –
>
> (e) the invention was described in (1) an application for patent, published under section 122(b), by another filed in the United States before the invention by the applicant for patent or (2) a patent granted on an application for patent by another filed in the United States before the invention by the applicant for patent, except that an international application filed under the treaty defined in section 351(a) shall have the effects for purposes of this subsection of an application filed in the United States only if the international application designated the United States and was published under Article 21(2) of such treaty in the English language.

6. Claims 1 to 3, 5 to 6, 8 to 11, 14 to 16, 18, 20 to 23, and 25 are rejected under 35 U.S.C. 102(e) as being anticipated by *Mayberry et al.*

Regarding independent claims 1 and 14, *Mayberry et al.* discloses a method and system for segmentation, extraction, summarization, and presentation of broadcast news, comprising:

"separating a multimedia data stream into audio, visual and text components" – at front end processor 150, data from streams in media source 102 are captured, including video imagery data 104, audio sample data 106, and closed captioned text data 108 (column 5, lines 46 to 63: Figure 1); Figure 1 shows the streams being separated for scene change detection 110, speaker change detection 116, and closed caption preprocessing 118;

"segmenting the audio, visual and text components of the multimedia data stream based on semantic differences, wherein frame-level features are extracted from the segmented audio component are in a plurality of subbands" – files representing imagery 104, audio 106, and closed captioned text 108 are fed to Broadcast News editor 100 to complete functions for segmentation and classification of news programs (column 5, lines 56 to 66: Figure 1); Broadcast News Editor provides for speaker change detection 116 and speech transcription 117, which involve elements of speech recognition; implicitly, speech recognition produces "frame-level features" from frames of speech ("the segmented audio component") in the form of cepstrals, which are obtained by Fourier transforms ("in a plurality of subbands");

"identifying at least one target speaker using the audio and visual components" – audio stream 106 is processed using speaker change detection 116, which identifies a speaker change event when the speaker-dependent components of the audio signal change more than a predetermined amount; frame classification 112 discovers single or double anchors (column 13, lines 27 to 39);

"identifying semantic boundaries of text for at least one of the identified target speakers to generate semantically coherent text blocks" – text stream event processing detects word patterns in closed captioned text to identify introductory phrases for story segments (column 10, line 54 to column 12, line 65); these phrases denote "semantic boundaries of text" for reporter to anchor hand-off or anchor to reporter hand-off ("for at least one of the identified target speakers");

"generating a summary of multimedia content based on the audio, visual and text components, the semantically coherent text blocks and the identified target speaker" – once individual news stories are identified, a story classifier 133 is used to identify a gist or theme for each story; the gist portion automatically consolidates large volumes of text, as provided by closed captioned stream 108 or converted audio stream 106 into a relevant summary; a database stores video theme records 312, video gist records 314, and story summary records 315 (column 15, line 61 to column 16, line 34);

"deriving a topic for each of the semantically coherent text blocks based on a set of topic category models" – a key frame is selected from a story segment to identify a speaker and a topic (column 17, lines 38 to 44);

"generating a multimedia description of the multimedia event based on the identified target speaker, the semantically coherent text blocks, the identified topic, and the generated summary" – Figures 16 to 19 show a story index having a multimedia description including a target speaker of "Leon" or "Joie", semantically coherent text blocks of "White House Oversight Board", an identified topic of "Clinton White House", and a summary for each record of "are in Brazzaville, Congo, ready to help evacuate Americans from neighboring Zaire, should that become necessary."

Regarding claims 2, 3, 6, 15, 16, and 18, *Mayberry et al.* discloses identifying multimedia content types including anchors and commercials of news broadcasts (column 13, line 52 to column 14, line 4).

Regarding claim 5, *Mayberry et al.* discloses words, images, or sounds in content of image frames, or visual features that serve as informative key frames (column 6, line 54 to column 7, line 2); these are elements of a video clip ("clip level features").

Regarding claims 8 and 20, *Mayberry et al.* discloses at least text descriptions and video descriptions (Figures 16 to 19).

Regarding claims 9 and 21, *Mayberry et al.* discloses storing files and summaries in a data base management system 140 (column 6, lines 4 to 9; column 16, lines 17 to 34: Figure 1).

Regarding claims 10, 11, 22, and 23, *Mayberry et al.* discloses presenting descriptions and summaries to a user pursuant to a search, and presentation to a user

by clicking on one of selected key frames for playback and viewing of a video clip (column 16, line 34 to column 18, line 29: Figures 15 to 19).

Regarding claim 25, *Mayberry et al.* discloses data are made available to a browser-enabled client 170 ("a terminal that displays") (column 6, lines 4 to 9: Figure 1).

Claim Rejections - 35 USC § 103

7. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

> (a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negatived by the manner in which the invention was made.

8. Claims 4, 7, 17, and 19 are rejected under 35 U.S.C. 103(a) as being unpatentable over *Mayberry et al.* in view of *Ahmad et al.*

Concerning claims 4 and 17, *Mayberry et al.* omits details of converting a multimedia data stream from an analog multimedia data stream to a digital multimedia data stream, and compressing the digital multimedia stream. However, it is well known that multimedia files generally originate as analog signals, and that MPEG provides for digitization and compression of multimedia files. *Ahmad et al.* teaches a browser for navigating audiovisual data of a body of information from an on-line news service or wire service (Abstract), where analog television signals must be digitized before being used in digital processing. This is accomplished with a conventional A/D conversion method and apparatus. Further, it is desirable to compress the data to increase the amount of data that can be stored in a storage device. Television can be compressed

according to MPEG, JPEG, or MJPEG video compression standards, and text data can be compressed using conventional text file compression programs, such as PKZIP. (Column 12, lines 3 to 28) Thus, the objectives are to digitize an analog television signal so as to enable digital processing, and to compress the data so as to decrease storage requirements. It would have been obvious to one having ordinary skill in the art to digitize and compress a multimedia file as taught by *Ahmad et al.* in an automated segmentation and summarization method and system for broadcast news of *Mayberry et al.* for purposes of enabling digital processing and decreasing storage requirements.

Concerning claims 7 and 19, *Mayberry et al.* omits a detail of speaker identification with Gaussian Mixture Models. However, it is well known to utilize Gaussian Mixture Models in speech recognition and speaker recognition because these represent more accurate models for recognizing speech. *Ahmad et al.* teaches a browser for navigating audiovisual data of a body of information from an on-line news service or wire service (Abstract), where Gaussian Mixture Models are employed for speaker recognition so as to partition audio data. (Column 24, Lines 10 to 29) It would have been obvious to one having ordinary skill in the art to utilize Gaussian Mixture Models as taught by *Ahmad et al.* in an automated segmentation and summarization method and system for broadcast news of *Mayberry et al.* for a purpose of partitioning audio data by voice recognition.

9. Claims 12, 13, and 24 are rejected under 35 U.S.C. 103(a) as being unpatentable over *Mayberry et al.* in view of *Wrench, Jr. et al.*

Mayberry et al. suggests that stories may be segmented by silence (column 13, line 63), but omits details of speaker recognition, where frame level features are extracted from at least three subbands, and frame level features in the three subbands are at least one of volume, zero crossing rate, pitch period, frequency centroid, frequency bandwidth and energy ratios. However, it is well known that speech recognition and speaker recognition involve obtaining frames of speech, extracting features in at least three subbands, and identifying features including at least a speaker's pitch in order to identify a speaker.

Wrench, Jr. et al. teaches a multiple parameter speaker recognition system, where front end processing involves a 512 point Fast Fourier Transform (FFT). (Column 8, Lines 44 to 59) Thus, there are 512 frequency subbands produced by the FFT, which is at least "three subbands". Then, extracted features are cepstral coefficients, and frame energy is compared to a threshold for identifying a speaker ("energy ratios"). (Column 8, Line 60 to Column 9, Line 30) Measurement of a zero crossing rate in several broad frequency bands to give an estimate of formant frequencies is another means of representing a speech signal for speaker recognition. (Column 1, Lines 37 to 49) The objective is to improve a method of speaker recognition by multiple parameters so that a speaker does not have to repeat a particular phrase. (Column 2, Lines 53 to 65) It would have been obvious to one having ordinary skill in the art to extract frame level features including at least energy ratios and zero crossings from at least three bands as taught by *Wrench, Jr. et al.* in an automated segmentation and summarization method and system for broadcast news of *Mayberry et al.* for a

purpose of improving speech recognition so that a speaker does not have to repeat a phrase.

Conclusion

10. The prior art made of record and not relied upon is considered pertinent to Applicants' disclosure.

Wilcox et al. and Gupta et al. disclose related art.

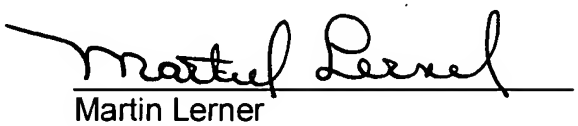
Any inquiry concerning this communication or earlier communications from the examiner should be directed to Martin Lerner whose telephone number is (571) 272-7608. The examiner can normally be reached on 8:30 AM to 6:00 PM Monday to Thursday.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, David R. Hudspeth can be reached on (571) 272-7843. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see http://pair-direct.uspto.gov. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free). If you would like assistance from a

USPTO Customer Service Representative or access to the automated information system, call 800-786-9199 (IN USA OR CANADA) or 571-272-1000.

ML
5/22/06



Martin Lerner
Examiner
Group Art Unit 2626